

June 1997

THE DESIGN AND ANALYSIS & SALMONID TAGGING STUDIES IN THE COLUMBIA BASIN

VOLUME II: ESTIMATING SALMONID SURVIVAL WITH COMBINED PIT-CWT

Technical Report



DOE/BP-35885-11



This report was funded by the Bonneville Power Administration (BPA), U.S. Department of Energy, as part of BPA's program to protect, mitigate, and enhance fish and wildlife affected by the development and operation of hydroelectric facilities on the Columbia River and its tributaries. The views of this report are the author's and do not necessarily represent the views of BPA.

This document should be cited as follows:

Newman, Ken - Division of Statistics, University of Idaho, The Design And Analysis & Salmonid Tagging Studies In The Columbia Basin, Volume II: Estimating Salmonid Survival With Combined PIT-CWT Tagging, Bonneville Power Administration, Contract No. 1987BP35885, Project No. 199105100, 28 electronic pages (BPA Report DOE/BP-35885-11)

This report and other BPA Fish and Wildlife Publications are available on the Internet at:

<http://www.efw.bpa.gov/cgi-bin/efw/FW/publications.cgi>

For other information on electronic documents or other printed media, contact or write to:

Bonneville Power Administration
Environment, Fish and Wildlife Division
P.O. Box 3621
905 N.E. 11th Avenue
Portland, OR 97208-3621

Please include title, author, and DOE/BP number in the request.

THE DESIGN AND ANALYSIS OF SALMONID TAGGING
STUDIES IN THE COLUMBIA BASIN

VOLUME II

ESTIMATING SALMONID SURVIVAL WITH COMBINED.
PIT-CWT TAGGING

Prepared by:

Ken Newman

Division of Statistics
University of Idaho

Prepared for:

U. S. Department of Energy
Bonneville Power Administration
Environment, Fish and Wildlife
P.O. Box 3621
Portland, OR 97208-362 1

Project Number **91-051-00**
Contract Number **DE-BI79-87BP35885**
Task Order **Number DE-AT79-91BP16570**

PREFACE

Project 91-051 was initiated in response to the Endangered Species Act (ESA) listings in the Snake River Basin of the Columbia River system. Primary objectives and management implications of this project include: (1) to address the need for further synthesis of historical tagging and other biological information to improve understanding and to help identify future research and analysis needs; (2) to assist in the development of improved monitoring capabilities, statistical methodologies, and software tools to assist management in optimizing operational and fish passage strategies to maximize the protection and survival of listed threatened and endangered Snake River salmon populations and other nonlisted stocks in the Columbia River Basin; and (3) to design better analysis tools for evaluation programs.

The following report addresses measure 5.0F.5 of the 1994 NPPC Fish and Wildlife Program with emphasis on improved design and analysis capabilities related to the conduct of **salmonid** tagging studies in the **Columbia River** Basin. This report addresses the need to study the fate of salmon smolt in-river and their subsequent return as adults. **Double-tagging** procedures are investigated where PIT-tags would be used to provide in-river survival data during **smolt** outmigrations and coded-wire tags (CWT) used to provide adult return information. This report provides statistical models for the analysis of the joint data as well as recommendations on optimal tagging studies. Study costs and stress on smolt can be reduced by only PIT-tagging a subset of all the fish coded-wire-tagged, while retaining the information content and sampling precision. This analysis illustrates how research goals can be achieved while efficiently using fiscal and natural resources.

Executive Summary

Passive Integrated Transponder (PIT) tags and Coded Wire Tags (CWTs) in combination can provide information about salmonid survival that single tag releases may not. The release and recapture protocol affects which survival and recapture rates can be estimated and the precision of the estimates. For the particular case of Columbia river salmonids tagged with both PIT tags and CWTs, three different release and recapture protocols were evaluated.

Protocol I : fish are scanned for PIT tags at all possible recapture sites, e.g., dams and fisheries, and tagged fish are released at each of these sites.

- **Benefits:** Survival and recapture rates for every interval between recapture points, except between the next-to-last and last point, can be estimated. For example, so long as two upstream recapture sites are present for returning adult salmonids, ocean harvest rate and ocean natural mortality rate can be separately estimated.
- **Costs:** Fishery catches must be scanned for PIT tags and tagged fish must be re-released from the fisheries (causing subsequent handling mortality).

Protocol II : fish are scanned for PIT tags at all possible recapture sites, but released only at dams, not in fisheries.

- **Benefits:** Survival and recapture rates for outmigrating juveniles can be estimated (and are identical to Protocol I), but caught fish are not re-released.
- **Costs:** Ocean fishery harvest rate cannot be separated from natural ocean mortality.

Protocol III : like Protocol II, but only CWT recoveries, not PIT tag readings, are available in the fisheries.

- **Benefits:** Exact same parameters can be estimated as for Protocol II without scanning caught fish for PIT tags.
- **Costs:** Parameter estimates are less precise than under Protocol II.

Protocol I is an idealised procedure that serves as a benchmark for comparison with the more feasible Protocols II and III.

Also examined was the precision of estimates under full and mixed tagging schemes which are defined as

Full tagging : all the fish in a group contain both a CWT and a PIT tag;

Mixed tagging : some of the fish receive both tags, while the remainder receive just a single tag (either a CWT or a PIT tag).

The loss in precision of mixed tagging compared to full tagging with the same total release numbers can be great. For example, based on simulations of a total release of 30,000 fish, when 90% of the fish got CWT-only and 10% of the fish got both tags, the loss of precision in estimates of survival rate was nearly three-fold relative to all getting both tags. Increasing the percentage of fish getting both tags beyond 40 or 50%, however, resulted in relatively small gains in precision. The issue of the optimal mixed tagging scheme for maximizing precision, given a fixed cost was also addressed.

Contents

1	Introduction	3
1.1	CWTs and PIT tags	3
1.2	Tag recapture points for Columbia River salmonids	3
1.3	Notation	4
1.4	Assumptions	4
2	Three release-recapture protocols	5
2.1	Protocol <i>I</i> : Complete capture history known and releases at all points of recapture	5
2.2	Protocol <i>II</i> : Complete capture history but no fisheries' releases	6
2.3	Protocol <i>III</i> : Incomplete capture histories and no fisheries' releases	8
2.4	Contrasting the recapture-release protocols	8
2.5	Special Case: $J_5 F_1 A_2 . II$	10
3	Mixed tagging schemes	10
3.1	Contrasting full and mixed tagging	12
4	Discussion	13
5	References	18
A	MLEs for Protocol <i>I</i>	17
B	MLEs for Protocol <i>II</i>	20
C	MLEs for Protocol <i>III</i>	21
D	Likelihoods for mixed tagging	22

1 Introduction

The aim of this report is to discuss estimation of fish survival and capture probabilities when coded wire tags (CWT) and passive integrated transponder (PIT) tags are used in combination. The focus of application is Columbia River salmonids, *Oncorhynchus spp.*, migrating past dams and through fisheries. The effect on estimable parameters of three different release-recapture protocols are compared:

Protocol I : fish are scanned for PIT tags at all possible recapture sites, e.g., dams and fisheries, and tagged fish are released at each of these sites.

Protocol II : fish are scanned for PIT tags at all possible recapture sites-but released only at dams, not in fisheries.

Protocol III : like Protocol II; but only CWT recoveries, not PIT tag readings, are available in the fisheries.

Also compared are full and mixed tagging schemes defined as

Full tagging : all the fish in a group receive both a CWT and a PIT tag;

Mixed tagging : some of the fish receive both tags, while the remainder receive just a single tag (either a CWT or a PIT tag).

1.1 CWTs and PIT tags

Coded wire tags and passive integrated transponder tags are two popular methods of tagging Pacific coast salmonids. CWTs are small round metal cylinders, about 1.1 mm in length and 0.25 mm in diameter (Nielsen, 1992), notched with a binary coding system. The number of possible codes is moderately restricted under this coding system and the codes are group specific, not individual fish specific. The CWT marking procedure for Pacific salmonids has been to insert the CWT into the head of a juvenile fish and to simultaneously remove the adipose fin. Upon recapture, recapture of the missing adipose fin indicates the likely presence of a CWT, the fish is sacrificed and the CWT is extracted and read. CWTs have been in use since the mid-1960s (Jefferts, Bergman, and Fiscus, 1963) and have been used to estimate survival rates and harvest rates as well as learn more about the geographic range of a fish stock. They are the primary data source for the development of Pacific ocean salmon management models used by state and federal agencies.

PIT tags are small plastic tubes roughly ten times larger than CWTs, 12 mm in length and 2.1 mm in diameter (Prentice, Flagg, and McCutcheon 1990; Nielsen, 1992). In the tube is a microprocessor and a metal antenna coil that when excited by a magnetic field causes the microprocessor to emit a radio signal. The microprocessor can code for a much greater range of identifiers than the binary system used by CWTs, so fish specific identifiers are used. PIT tags are inserted into the lower body cavity of juvenile salmonids and are read when the fish passes through tubes, located in dam bypasses, that generate a magnetic field. Handheld scanners can be used to read the tags as well. PIT tags have been in use since the mid-1980s (Prentice, et al. 1990) and can provide more detailed information than CWTs, since individual fish can be identified without sacrificing the fish and the location of a fish at multiple points in time can be made. For more details on CWTs and PIT tags see Nielsen (1992).

1.2 Tag-recapture points for Columbia River salmonids

Juvenile outmigration for Columbia River system salmonids begins with departure from their natal stream or hatchery, and movement downstream to the ocean, most travelling through one or more dams. Travel through the dams follows one of four routes, through juvenile fish bypass systems, through the turbines, over dam spillways, or through the lock system. The bypass systems on several of the dams contain magnetic coils that 'read' PIT tags in fish, thus providing a means of identifying some of the fish that reach the

dams: Fish surviving the river will spend one to four years in the ocean where some are caught by fisheries¹. Current sampling programs of the ocean fishery catch include **examination** for missing adipose-fins and lead to recapture of some of the CWT tagged fish, but scanning for PIT tags is not done. Maturing fish that survive the ocean fishery fish return to the Columbia River where some may be caught by in-river fisheries, which again are sampled for **CWTs** but not **PIT** tags. Currently the sole PIT tag detector in the Columbia system for returning fish is located in the fish ladder at Lower Granite Dam **located** 173 km upstream of the confluence of the Columbia and Snake Rivers.

1.3 Notation

Recapture scenarios will be denoted by $J_i F_j A_k$ when there are i juvenile recapture points, j fisheries, and k returning adult recapture points. Detection points are viewed sequentially, i.e., juvenile recapture points precede fisheries which precede adult recapture points, there is no spatial component, and there is no practical distinction made between recovery in an ocean fishery and an in-river fishery. In the context of Columbia river salmonids, juvenile recapture points will be dams encountered by outmigrating salmon. To distinguish different release-recapture protocols, a Roman numeral will be appended to the capture scenario, e.g., $J_2 F_1 A_1.II$.

The remaining notation will largely follow that of **Burnham, Anderson, White, Brownie, and Pollock** (1987).

R_i = Number of marked fish released at location i

S_i = Conditional survival probability from location i to location $i + 1$, given that fish is alive at location i

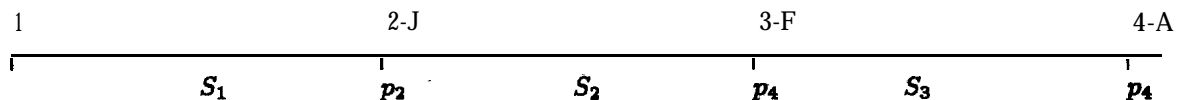
p_i = Detection probability at location i

q_i = $1 - p_i$

m_{ij} = Number of fish recaptured for the first time at location j after release from location i ($j > i$)

r_i = Total recaptures of fish released from location i .

A sketch of the release and recapture points and survival and capture probabilities is given below for scenario $J_1 F_1 A_1$, a situation with the initial release location (1), one juvenile recapture location (2), one fishery (3), and one adult recapture location (4).



1.4 Assumptions

The baseline assumptions follow many of those listed by **Burnham, et al.** (1987);

1. Marked fish are representative of the population of fish about which one wishes to **estimate** survival and recapture probabilities.
2. Numbers of marked fish released are known exactly.
3. Marks are identified without error.
4. Release and recapture times occur in brief intervals.
5. The fate of each marked fish is independent of other marked fish.
6. Captured and re-released fish have the same subsequent survival and recapture rates as do uncaptured fish.

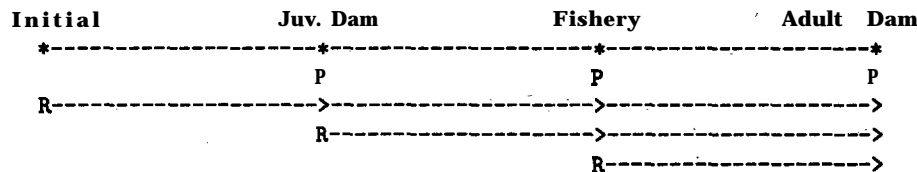
The fall race of chinook salmon, *Oncorhynchus tshawytscha*, are more vulnerable to ocean fishing mortality than are steelhead *Oncorhynchus mykiss* and spring chinook salmon.

2 Three release-recapture protocols

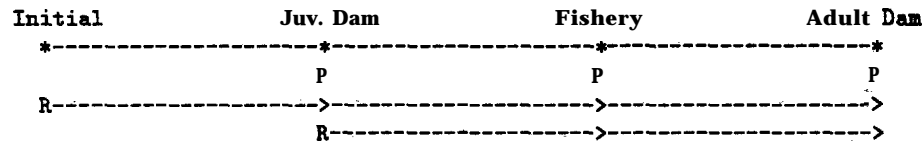
For all three release-recapture protocols discussed in this section, assume that fish receive both a PIT tag and a CWT, and the adipose fin is removed. Under Protocols I and II **CWTs** are technically unnecessary assuming fish at any recapture point will be scanned for PIT tags. As a practical means of avoiding scanning all fish in a fishery sample, however, the missing adipose fin may serve as a convenient marker, and current tagging practices reserve the missing adipose fin for fish containing a CWT.

For later reference a schematic diagram contrasting the release and recapture steps for each protocol under the simple case of one juvenile dam, one fishery, and one adult dam, $J_1F_1A_1$, is drawn below. The letters R, P, and C indicate that Releases, PIT tag scanning, and CWT reading, respectively, took place at the location site:

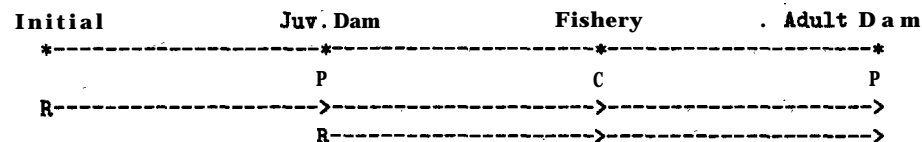
Protocol I



Protocol II



Protocol III



2.1 Protocol I: Complete capture history known and releases at all points of recapture

This protocol is the most labor expensive of the three, but provides the most information about survival and capture rates. At every possible capture point fish are examined for PIT tags and at least some recaptures are m-released. If the capture point is an ocean fishery, having fishermen release tagged fish would be novel, but there is some precedence for this in troll fisheries releasing undersize fish. One statistical concern is the effect of capture and handling on subsequent mortality rates of released fish (see Assumption 6 above).

Cormack (1964) proposed a probability model for such a situation and derived the maximum likelihood estimates (**MLEs**) for S_i and p_i and corresponding variance estimates². I will review the Cormack model from the release-recapture viewpoint described by Brownie, Anderson, Burnham, and Robson (1985) and for simplicity will look at the model, $J_1F_1A_1.I$.

Making the six assumptions in Section 1.4, the probability model for recaptures is based on the multinomial distribution. The perspective is that each group of fish released from a recapture site is a new release group and the locations of first recapture are the essential data provided by that group. If a fish is recaptured and m-released, it now belongs to a new release group and its next recapture is the essential datum for that new group; Therefore additions of marked fish at any recapture point as well as deliberate deletions (such as those put onto barges) are easily dealt with. Convenient summaries of this model are the (reduced) m-array,

²It is also a special case of the Jolly-Seber model (Jolly 1965, Seber 1965), the restriction in this case is that unmarked fish are not counted and total population size cannot be estimated.

m_{ij} , the recoveries at site j from release site i (Burnham, et al. 1987) and the corresponding conditional probabilities (see Table 1). The likelihood function for $J_1 F_1 A_1 . I$ is the product of the likelihood for the first release group times the conditional likelihoods of subsequent releases:

$$L(S_1, p_2, S_2, p_3, S_3, p_4) \propto (S_1 p_2)^{m_{12}} (S_1 q_2 S_2 p_3)^{m_{13}} (S_1 q_2 S_2 q_3 S_3 p_4)^{m_{14}} \times \\ (1 - S_1 p_2 - S_1 q_2 S_2 p_3 - S_1 q_2 S_2 q_3 S_3 p_4)^{R_1 - r_1} \times \\ (S_2 p_3)^{m_{23}} (S_2 q_3 S_3 p_4)^{m_{24}} (1 - S_2 p_3 - S_2 q_3 S_3 p_4)^{R_2 - r_2} \times \\ (S_3 p_4)^{m_{34}} (1 - S_3 p_4)^{R_3 - r_3} \quad (1)$$

Not all the parameters are estimable, namely S_3 and p_4 ; in equation (1) the two parameters are always paired and cannot be separated. For any release-recapture protocol the survival rate to the last point of recapture is not separable from the recapture rate at that point. For example, to separate ocean survival of Columbia River releases from ocean recapture rates (harvest rates), two points of recapture beyond the ocean are necessary.

The maximum likelihood estimates for the estimable parameters in $J_1 F_1 A_1 . I$ are

$$\hat{S}_1 = \frac{r_2 m_{12} + R_2(r_1 - m_{12})}{R_1 r_2} \quad (2)$$

$$\hat{p}_2 = \frac{r_2 m_{12}}{r_2 m_{12} + R_2(r_1 - m_{12})} \quad (3)$$

$$\hat{S}_2 = \frac{r_2[(m_{13} + m_{23})m_{34} + R_3(m_{14} + m_{24})]}{R_2 m_{34}(m_{13} + m_{14} + r_2)} \quad (4)$$

$$\hat{p}_3 = \frac{(m_{13} + m_{23})m_{34}}{(m_{13} + m_{23})m_{34} + R_3(m_{14} + m_{24})} \quad (5)$$

$$\widehat{S_3 p_4} = \frac{m_{34}}{R_3} \quad (6)$$

The general solution to maximum likelihood estimates for $J_i F_j A_k . I$ was given by Cormack (1964). Appendix A contains an alternative expression for the general solution along with variance estimates for \hat{S} and \hat{p} .

2.2 Protocol II: Complete capture history but no fisheries' releases

Under this protocol adipose-clipped fish are scanned in the fishery catch for a PIT tag, but fish recovered in a fishery are not released, For $i = j = k = 1$ the m-array and corresponding conditional probabilities are shown in Table 1³.

The estimable parameters are $S_1, p_2, S_2 p_3$, and $S_2 q_3 S_3 p_4$ with MLEs:

$$\hat{S}_1 = \frac{r_2 m_{12} + R_2(r_1 - m_{12})}{R_1 r_2} \quad (7)$$

$$\hat{p}_2 = \frac{r_2 m_{12}}{r_2 m_{12} + R_2(r_1 - m_{12})} \quad (8)$$

$$\widehat{S_2 p_3} = \frac{r_2}{R_2} \left(\frac{m_{13} + m_{23}}{m_{13} + m_{14} + m_{23} + m_{24}} \right) \quad (9)$$

$$S_2 q_3 \widehat{S_3 p_4} = \frac{r_2}{R_2} \left(\frac{m_{14} + m_{24}}{m_{13} + m_{14} + m_{23} + m_{24}} \right) \quad (10)$$

The survival rate between the juvenile dam and the fishery, S_2 , cannot be estimated separately from the fishery capture rate (contrast equations (4) and (5) with equation (9)). When re-releases are not made at

³This situation is identical to the setting described by Skalski in Appendix B of Dauble, Skalski, Hoffman, and Giorgi (1993). Juvenile fish were released above one dam, recaptured and possibly released at the dam, survivors then went through two more dams, but there were no releases at the second and third dams. Here the fishery corresponds to the second dam, and the dam encountered by returning adults corresponds to the third dam encountered by juveniles.

Table 1: m-array and conditional probabilities for the three protocol, $J_1 F_1 A_1$. ($q_i = 1 - p_i$.)

Protocol I m-array					
Release site	Release No.	Recapture Site			Total Recoveries
		Juv Dam	Fishery	Adult Dam	
Above Juv Dam	R_1	m_{12}	m_{13}	m_{14}	r_1
Juvenile Dam	R_2		m_{23}	m_{24}	r_2
Fishery	R_3			m_{34}	r_3

Protocol I conditional probabilities				
Release site	Recapture Site			
	Juv Dam	Fishery	Adult Dam	
Above Juv Dam	$S_1 p_2$	$S_1 q_2 S_2 p_3$	$S_1 q_2 S_2 q_3 S_3 p_4$	
Juvenile Dam		$S_2 p_3$	$S_2 q_3 S_3 p_4$	
Fishery			$S_3 p_4$	

Protocol II m-array					
Release site	Release No.	Recapture Site			Total Recoveries
		Juv Dam	Fishery	Adult Dam	
Above Juv Dam	R_1	m_{12}	m_{13}	m_{14}	r_1
Juvenile Dam,	R_2		m_{23}	m_{24}	r_2

Protocol II conditional probabilities				
Release site	Recapture Site			
	Juv Dam	F i s h e r y	Adult Dam	
Above Juv Dam	$S_1 p_2$	$S_1 q_2 S_2 p_3$	$S_1 q_2 S_2 q_3 S_3 p_4$	
Juvenile Dam		$S_2 p_3$	$S_2 q_3 S_3 p_4$	

Protocol III m-array					
Release site	Release No.	Recapture Site			Total Recoveries
		Juv Dam	Fishery	Adult Dam	
Above Juv Dam	R_1	m_{12}	m_{13}	m_{14}	?
Juvenile Dam	R_2			m_{24}	?

Protocol III conditional probabilities				
Release site	Recapture Site			
	Juv Dam	Fishery	Adult Dam	
Above Juv Dam	$S_1 p_2$	Σ	$S_1 q_2 S_2 q_3 S_3 p_4$	
Juvenile Dam			$S_2 q_3 S_3 p_4$	

a point in the capture sequence, the survival rate to that point cannot be separated from capture rate at that point.

In general, estimates of S_i and p_i for all intervals up to the last juvenile dam are estimable and are identical to the parameter estimates under *JFA.I* (e.g., equations (2) and (3) equal (7) and (8)). Consequently the variance estimates for those parameters are identical, too. Extensions in terms of additional juvenile dams and/or adult dams are discussed in Appendix B.

2.3 Protocol *III*: Incomplete capture histories and no fisheries' releases

Under the third protocol harvested fish are not re-released; they are checked for CWTs, but not PIT tags. Thus the complete capture history for fish recaptured in the fishery is unknown: recoveries in the fishery cannot be distinguished by those previously recaptured and not recaptured at the juvenile dam. The m -array and conditional probabilities for $J_1F_1A_1.III$ are shown in Table 1. $m_{.3}$, e.g., the sum of m_{13} and m_{23} , is known, but the individual components are not. This is the current recovery situation for commercial and recreational fisheries. The conditional probability for $m_{.3}$ (denoted \sum) is a complicated summation (details in Appendix C).

The same parameters as for Protocol *II* can be estimated, e.g., S_1 , p_a , S_2p_3 , and $S_2q_3S_3p_4$ for $J_1F_1A_1$, but the parameter estimation is more technically involved. An approach to computing MLEs is an approximation to the EM (Expectation and Maximization) algorithm (Dempster, Laird, and Rubin, 1977). The EM algorithm proceeds by taking initial estimates of the parameters and using, them to partition $m_{.3}$ into estimates of m_{13} and m_{23} . The estimated values of m_{13} and m_{23} are then used to re-estimate the parameters exactly as if it were a Protocol *II* situation- (Appendix C).

Limited extensions to Protocol *III* beyond $i = j = k = 1$ have been made, but analytical variance calculations have yet to be worked out (Appendix C). An example of the estimation procedure and a contrast with Protocol *II* precision is given in Section 2.4.

2.4 Contrasting the recapture-release protocols

The setting described by Protocol *I* is best in terms of estimating the most parameters possible with the least variance. On the other hand, it is the most costly program, the most logistically difficult, and perhaps the most prone to violation of assumptions because of subsequent handling mortality. As mentioned above, estimates of juvenile survival and recapture rates are identical for Protocols *I* and *II*. If there is a significant fishery, the primary advantage of Protocol *I* over Protocol *II* is the ability to separately estimate the fishery harvest rate and the ocean survival rate between the last juvenile dam and the fishery. For some Columbia River salmonids distinguishing these two parameters is vital to understanding the effects of harvest management plans. For instance a combined survival rate to the fishery and the harvest rate of 0.1% could be due, for example, to 1% survival and 10% harvest rate on the survivors, or a 0.5% harvest rate and 20% harvest rate.

The situations described by Protocols *II* and *III* do not differ in terms of estimability, both can estimate the same parameters, but estimates from Protocol *III* will be less precise. Approximate variances can be calculated easily for Protocol *II*, but analytical estimates of the variances under Protocol *III* are extremely complicated. As an alternative to analytical comparisons, a simulation study was done for the simplest case of one juvenile dam, one fishery, and one adult dam. Factors that influence the precision of estimates include initial release size, R_1 , and release size at the juvenile dam, R_2 , and the parameter values themselves. Only the release sizes and S_1 were varied in the simulation, where R_1 was set equal to (30,000, 40,000, 50,000), S_1 set equal to (0.9, 0.8, 0.7), and R_2 was the product of $R_1S_1p_2$. The release size at the juvenile dam then was the expected number of fish surviving to and being recaptured at the dam. The remaining parameter values were fixed at the values shown in Table 2.

Figure 1 shows the parameter estimates for the first 50 simulations with $R_1=40,000$ and $S_1=0.8$. The horizontal lines going through the plots mark the true values. The estimates based on Protocol *III* are marked with a 3, and 2 for Protocol *II*. For S_1 and p_2 , the Protocol *III* estimates are much more variable than the Protocol *II* estimates; both sets of estimates- are accurate, but the variation of Protocol *III*

Table 2: Parameter values for simulation of Protocols *II* and *III* ($J_1 F_1 A_1$).

p_a	9.9	Detection rate at juvenile dam
S_2	0.01	Survival rate between juvenile dam and fishery
p_3	0.2	Harvest rate of fishery
S_3	0.2	Survival rate between fishery and adult dam
p_4	0.9	Detection rate at adult dam

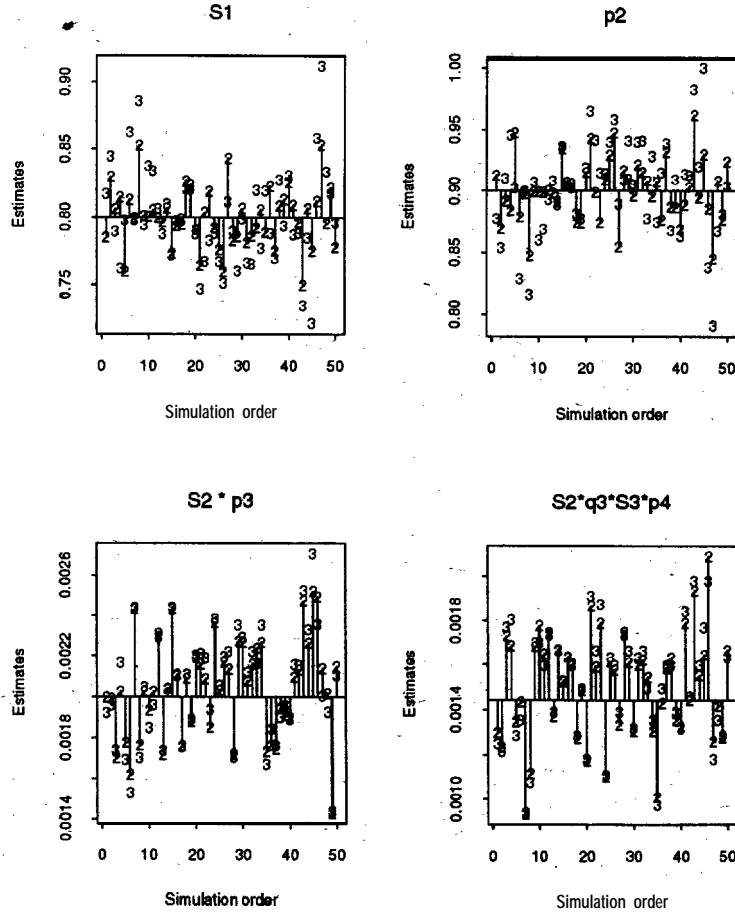


Figure 1: 50 simulated estimates of the parameters for Protocol *II* and Protocol *III* when $R_1=40,000$, $S_1=0.8$, $p_2=0.9$, $S_2=0.01$, $p_3=0.2$, $S_3=0.2$, and $p_4=0.9$. (2 corresponds to II and is connected by line to true value; 3 corresponds to III.)

Table 3: Percent Increase of Standard Errors of \hat{S}_1 for Protocol *III* relative to Protocol *II* based on 500 simulations.

S_1	R_1		
	30,000	40,000	50,000
0.9	62	59	63
0.8	72	58	42
0.7	56	58	61

estimates is consistently greater. Table 3 shows the relative increase in the standard errors of \hat{S}_1 based on 500 simulations under each release and S_1 combination. The loss in precision in estimating S_1 due to not scanning the fishery catch for PIT tags ranged from 42% to 72%. There does not seem to be a pattern to the increase in relative standard error and release numbers and S_1 over the range tried, but 500 simulations may be not be enough to provide stable estimates of the standard errors. The resulting ratios of standard errors of estimates of p_2 were nearly identical to those for S_1 . For the combined parameters S_2p_3 and $S_2q_3S_3p_4$, the loss in precision was minor, between 2 and 5%.

2.5 Special Case: $J_5F_1A_2.II$

For the Snake River salmonids a total of five outmigrating juvenile and two returning adult recapture facilities may eventually be operational. This special case was simulated assuming m-release capabilities at the juvenile and returning adult facilities, ocean or in-river harvest, but no release from fisheries.

The points of recapture for outmigration, in order of encounter on the Snake River and Columbia River, are Lower Granite Dam, Little Goose Dam, Lower Monumental Dam, McNary Dam, and Bonneville Dam. Then a 'single fishery takes place, and survivors return with possible detection at a Bonneville Dam adult facility and the Lower Granite adult facility. With survival rates to and recapture rates at each point, there are 16 parameters. Under Protocol *I* all but the last survival and recapture rates can be estimated, but the combined parameter S_8p_9 can be estimated. Under Protocols *II* and *III*, ten of the individual survival and recapture rate parameters and four other combinations of parameters can be estimated, where the first ten parameters are all the survival and recapture rates in the Snake and Columbia Rivers during outmigration. Exact maximum likelihood estimates can be made for the 14 parameter combinations.

As a demonstration of the precision of estimates, a simulation of a release of 30,006 fish with a full tagging strategy, namely both CWTs and PIT tags, was made under release-recapture Protocol II. Table 4 lists the parameters, their true values, and estimated values. For this range of parameter values and release size, the estimates were quite good.

3 Mixed tagging schemes

Three different mixed tagging schemes are considered. In the first case one subgroup gets both tags while the remainder get CWTs alone. In the second case one subgroup gets both tags and the remainder get PIT tags alone. In the third case the release group one subgroup gets both tags, a second subgroup gets PIT tags alone, and the remainder get CWTs alone. When only the PIT tag is inserted, the adipose fin is assumed left intact.

For any given release-recapture protocol, the same parameters estimable under full tagging are still estimable under mixed tagging, because a subgroup receiving both tags exists. The likelihoods for any given protocol under mixed tagging schemes are products of the likelihood for the combination group and the single tag only groups (see Appendix D). What is lost is precision relative to what would have been achieved if all fish had received both tags. What is gained is a reduction in tagging and tag reading costs.

To simplify comparisons, $i = j = k = 1$, namely, one juvenile dam, one fishery, and one adult dam, is

Table 4: Simulation of $J_5 F_1 A_2 II$ and parameter estimates

Parameters		True	MLEs
S_1	survival to L. Granite	0.700	0.701
p_2	recapture at L. Granite	0.800	0.798
S_2	recapture survival to L. Goose	0.800	0.800
p_3	recapture at L. Goose	0.900	0.901
S_3	survival to Lo Monumental	0.800	0.802
p_4	recapture at Lo Monumental	0.900	0.899
S_4	survival to McNary	0.750	0.747
p_5	recapture at McNary	0.800	0.799
S_5	survival to Bonneville	0.700	0.693
p_6	recapture at Bonneville	0.900	0.905
$S_6 p_7$	Ocean survival+harvest	0.030	0.032
$S_6 q_7 S_7 p_8$		0.033	0.036
$S_6 q_7 S_7 q_8$		0.002	0.002
$S_8 p_9$	survival and recapture to Bonnev.	0.828	0.815

Table 5: m-arrays for partial tagging situations under $J_1 F_1 A_1$

CWT alone				
Release site	Release No. ,	Recapture Site		Total
		Juv Dam	Fishery Adult Dam	Recoveries
Above Juv Dam	R_1	?	m . 3	?

PIT alone				
Release site	Release No. ,	Recapture Site		Total
		Juv Dam	Fishery Adult Dam	Recoveries
Above Juv Dam	R_1	m_{12}	?	m_{14}
Juvenile Dam	R_2	?	m_{24}	r_2

assumed in each case.

Case 1: Both and CWT-oily subgroups

For a CWT-only group no fish would be recaptured at either dam, nor would there be any reason to release CWT-only fish from the fishery, because none would be recaptured at the adult dam. The reduced m-array is very sparse in this case (Table 5). All that is known are the original release-size R_1 and the total number of CWT recoveries in the fishery. Since there is no means of externally distinguishing CWT-only fish from fish with both tags, the CWT-only fish would be scanned for PIT tags as well.

Case 2: Both and PIT tag-only subgroups

For a PIT tag-only release group no fish would be recaptured in a fishery assuming it is the missing adipose fin (that accompanies a CWT) that tips off the fisherman to the presence of a tag. Relatedly there could be no releases of PIT tag-only fish from the fishery. Therefore, the relevant data is the same for all three release-recapture protocol. The m-array for PIT tag-only fish is shown in Table 5. To estimate the parameters the EM algorithm can be used (see Appendix D).

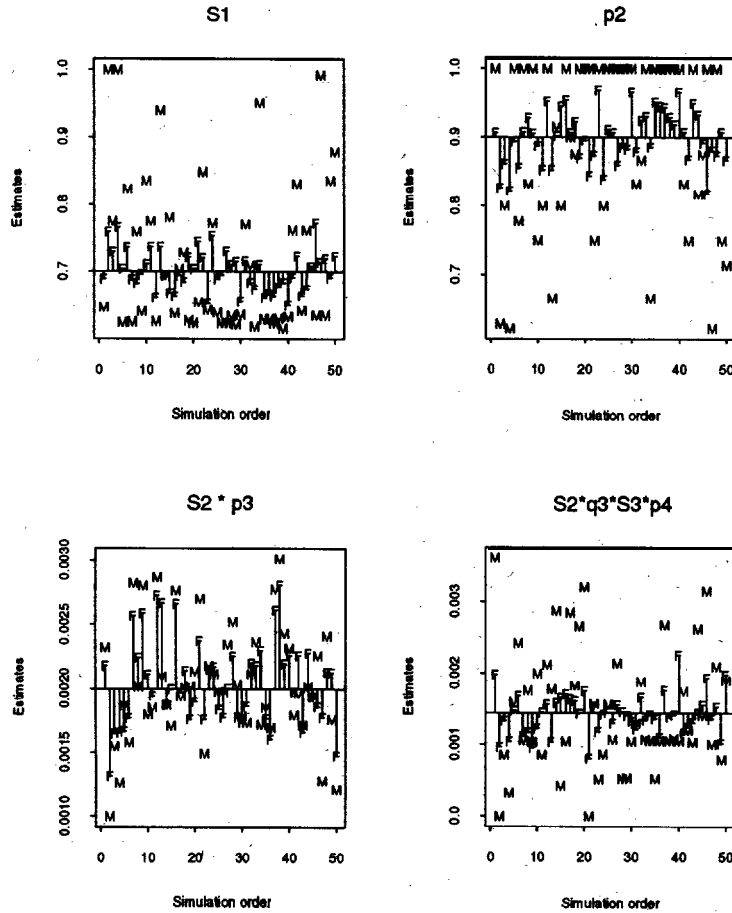


Figure 2: Comparison of 50 simulated estimates of the parameters for Protocol II based on full and mixed tagging (Case 1). (F corresponds to full and M to mixed)

Case 3: Both, PIT tag-only, and CWT-only subgroups

The features of cases 1 and 2 are simply combined. There are three different m-arrays, one for each group, each shown previously and the likelihood is the product of the three individual likelihoods.

3.1 Contrasting full and mixed tagging

The effect of mixed tagging on the quality of the parameter estimates was assessed for Case 1— namely, one subgroup received both tags and the remainder received CWTs alone. The release-recapture setting and protocol was $J_1 F_1 A_1 . II$. In all cases a total of 30,000 fish were released, S_1 was set equal to 0.7 and the remaining parameter values were again those in Table 2.

In one simulation the combined tagging rate was 10%. 3,000 fish were released with both tags (R_1) and all recaptured survivors were re-released at the first recovery site (R_2). The remaining 27,000 fish received CWTs alone (R_C). Figure 2 plots the first 50 estimates for the estimable parameter combinations for the full and mixed cases. The estimates were equally accurate, but the imprecision in the mixed tagging group was considerably greater. Table 6 compares the means and standard deviations based on 100 runs. The loss

Table 6: Means and Standard Errors for full and mixed tagging (Case 1) under Protocol II based on 100 simulations. **30,000 fish** were released in both cases. For the mixed tagging 27,000 received **CWTs** alone and the remaining 3,000 received both tags.

Parameter	True Value	Mean Full	Mean Mixed	SE Full	SE Mixed	% Increase in SE (Mixed:Full)
S_1	0.70	0.70	0.72	0.031	0.115	272%
p_2	0.90	0.90	0.89	0.039	0.127	226%
S_2p_3	0.062	0.0020	0.0020	0.00034	0.00045	32%
$S_2q_3S_3p_4$	0.0014	0.0014	0.0014	0.00029	0.00086	196%

of precision of mixed tagging relative to **full tagging** ranged from 42% to 222%.

The above simulation was repeated over a range of combined tagging rates of 10% to 90%. Figure 3 plots the empirical standard errors of the four estimable parameters. The F's mark the standard errors under full tagging and M's do the same for mixed tagging. Note that the **rate of** increase in precision as tagging fraction increases depends on the parameter being estimated. For instance, the change in the precision of estimates of the first survival rate, S_1 , is much greater going from 20 to 40% than is the change for the combined second survival rate and recapture rate S_2p_3 . Thus the relative importance of the parameters is a factor in determining an adequate tagging ratio with mixed tagging schemes. Focusing on the parameter S_1 , a tagging more than 40% of the release with both tags will give roughly equivalent precision to the full tagging strategy.

A related problem is to find the tagging strategy that minimizes variances for a **fixed** cost. Suppose that a PIT-CWT combination costs \$3 compared to **\$1** for a single CWT and that the total cost is fixed at \$90,009 Using the same parameter values as before under $J_1F_1A_1.II$ and Case 1, 100 simulations were carried out for five different combination-CWT only mixtures; each with a total cost of **\$90,000**. Figure 4 plots the resulting standard errors for each combination. As before, the optimal mix varies between parameters. For estimating S_1 and p_2 about 24,000 fish with both tags (and subsequently 18,000 fish with only **CWTs**) appears best. For estimating S_2p_3 , however, 18,000 fish with both tags (subsequently, 36,000 fish with only **CWTs**) is optimal.

4 Discussion

Conclusions that can be drawn from this work are the following:

1. **Re-releasing** captured fish from a fishery is necessary to separately estimate fishery harvest rates (the 'detection' rates) and survival from the fishery to the next recapture point.
2. Assuming fish are tagged with both PIT tags and **CWTs**, scanning fishery catches for **PIT** tags, as opposed to the the current practice of just looking for **CWTs**, can result in **sizeable gains** in precision.
3. Mixed tagging strategies, with rates for combination tags as **low** as **50%**, can provide **nearly** equal precision to full tagging strategies.

Some of the oversimplifications made herein point to areas for future work:

- Detection in the fishery catches has been oversimplified in two ways. First, catches are **generally** subsampled for tags, - some commercial catches **and** **sport** charter boat catches are randomly port sampled (with boats being the sampling units). If the sampling rate is known, this additional parameter can be separated from the ocean survival and harvest rate parameters. Second, fisheries do not occur in a single location at a single point in time, nor are all the fish present in a given location at a single point in time. The notion of an ocean, pm-harvest survival rate is therefore misleading. There are separate probabilities of survival to a given **place** at a given time. Likewise there is no single fishery

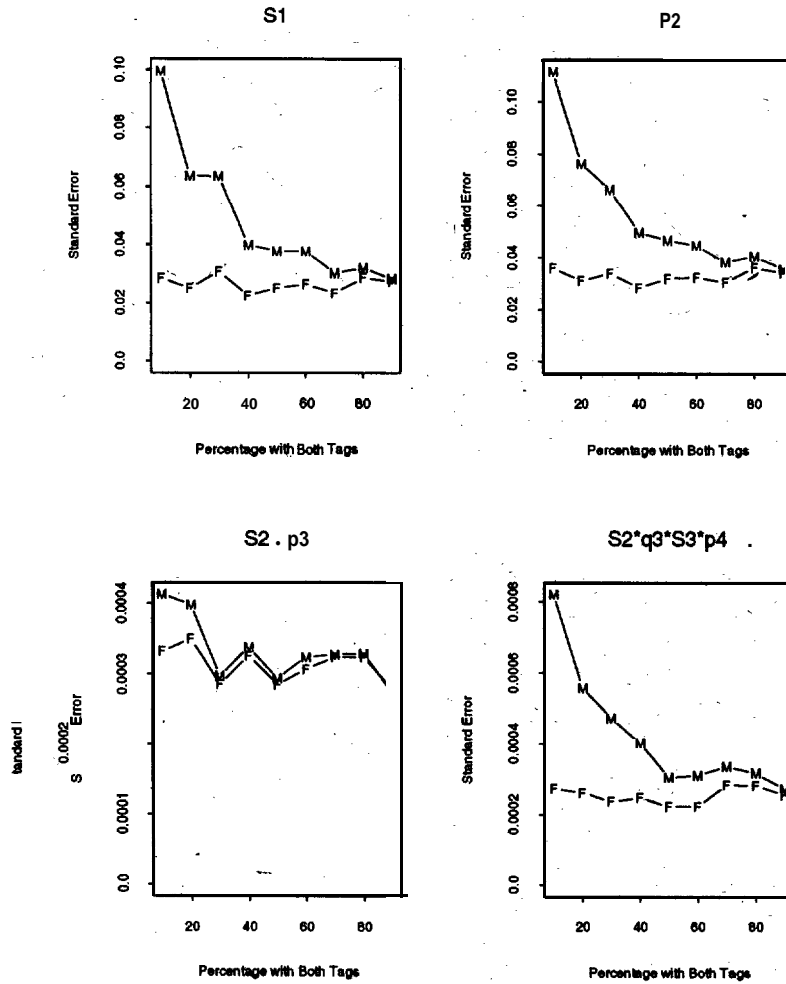


Figure 3: Standard errors as function of percentage of fish getting both tags under Protocol II based on Case 1. (F corresponds to full and M to mixed)

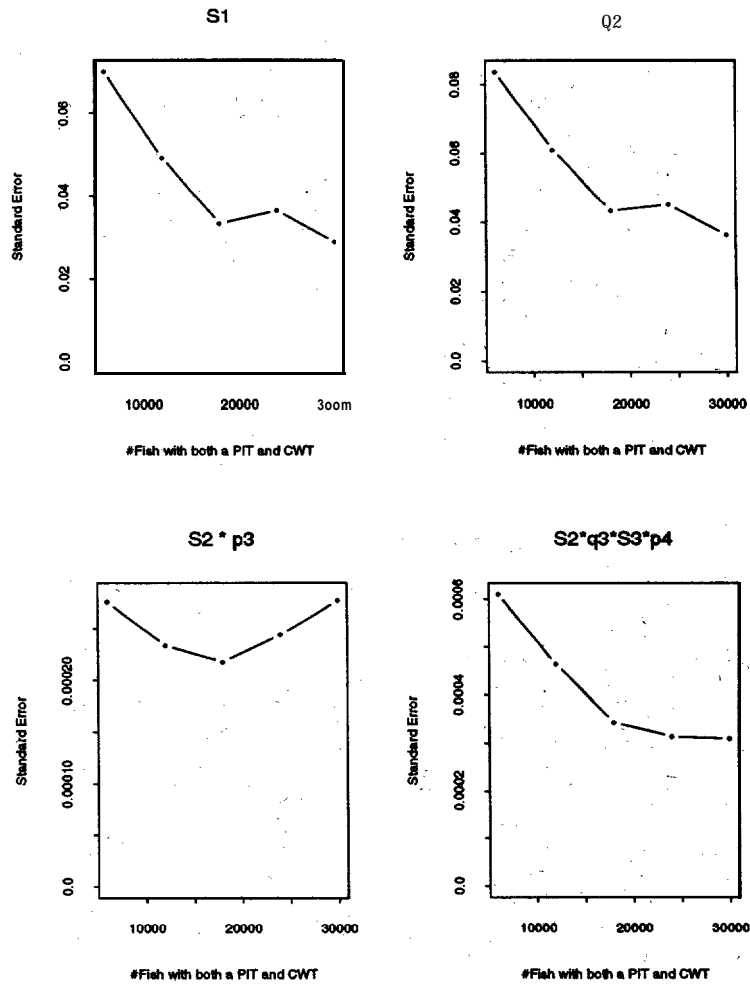


Figure 4: Standard errors as function of number of fish getting both tags under Protocol II -based on Case 1 when using both tags is 3 times as costly as CWT-only tagging.

separate probabilities of survival to a given place at a given time. Likewise there is no single fishery harvest rate, rather an array of probabilities distributed in space and time is more accurate. Further examination of the potential problems in the naive estimates is necessary.

- Differing maturation rates have been ignored. With cohorts maturing at different age classes, as is the case with chinook salmon, survival to the adult dams is a function of ocean survival rates, harvest rates, and maturation probabilities. Models would need to be extended to deal with such species and estimation of ocean survival and fishery rates will be confounded by the maturation schedule.
- Analytical estimates of the variances are preferable to simulations, for comparing protocols and tagging mixtures. For instance, given analytical formula for variance estimates and estimates of tagging and recovery costs, the optimal mix of tag subgroups could be determined more precisely and efficiently than the simulation method presented here. Relatedly, noting the apparent lack of a single optimal tagging mixture for estimating several parameters, an optimal mixture may be one that maximizes a measure of overall statistical information.

5 References

- Brownie, C., Anderson, D.R., Burnham, K.P., and Robson, D.S. (1985). *Statistical inference from band recovery data- a handbook*, U.S. Fish and Wildlife Service Resource Publication 156.
- Burnham, K.P., Anderson, D.R., White, G.C., Brownie, C. and Pollock, K.H. (1987). *Design and analysis methods for fish survival experiments based on release-recapture*, American Fisheries Society Monograph 5.
- Cormack, R. (1964). 'Estimates of survival from the sighting of marked animals', *Biometrika*, 51, 429-438.
- Dauble, D.D., Skalski, J., Hoffman, A.; and Giorgi A. E. (1993). 'Evaluation and application of statistical methods for estimating smolt survival', *Bonneville Power Administration*, Portland, OR.
- Davidson R.R., and Solomon, D.L. (1974). 'Moment-type estimation in the exponential family', *Communications in Statistics*, 3(11), 1101-1108.
- Dempster, A.P., Laird, N.M., and Rubin, D.B. (1977). 'Maximum likelihood from incomplete data via the EM algorithm', *Journal of the Royal Statistical Society, B*, 1, 1-38.
- Jefferts, K.B., Bergman, P.K., and Fiscus, H.F. (1963). 'A coded wire identification system for macro-organisms', *Nature*, 198, 460-462.
- Jolly, G.M. (1965). 'Explicit estimates from capture-recapture data with both death and immigration-stochastic models', *Biometrika*, 52, 225-247.
- Neilsen, L.A. (1992). *Methods of Marking Fish and Shellfish*, American Fisheries Society Special Publication 23.
- Prentice, E.F., Flagg, T.A., and McCutcheon, C.S. (1990). 'Feasibility of using implantable passive integrated transponder (PIT) tags in salmonids' in *Fish Marking Techniques* ed. by Parker, N.C., A.E. Giorgi, R.C. Heidinger, D.B. Jester, Jr., E.D. Prince, and G.A. Winans. American Fisheries Society Symposium 7, 317-322.
- Seber, G.A.F. (1965). 'A note on the multiple-recapture census', *Biometrika*, 52, 249-259.
- Tanner, M.A. (1996). *Tools for statistical inference. 3rd Ed.* Springer, New York.

Acknowledgements

I thank John Skalski for suggesting the idea of combining PIT tags and CWTs.

A MLEs for Protocol I

Cormack (1964) derived MLEs for Protocol I for arbitrary numbers of release-recapture points. I re-derive his results using the ideas of Davidson and Solomon (1974). This approach will be of value with some of the other release-recapture protocols. First I begin with the derivation for the special case of $J_1 F_1 A_1 . I$ and show that the method of moments estimators (MMEs) based on the minimal sufficient statistics are identical to the MLEs.

A convenient reparametrization, employed by Cormack (1964) and Seber (1965), is to let

$$\begin{aligned}\alpha_i &= S_i p_{i+1} \\ \beta_i &= S_i q_{i+1}\end{aligned}$$

This parametrization is 1:1 with S and p because

$$\begin{aligned}S_i &= \alpha_i + \beta_i \\ p_{i+1} &= \frac{\alpha_i}{\alpha_i + \beta_i}\end{aligned}$$

The likelihood for the release-recapture model can be re-written as

$$L \propto \alpha_1^{m_{12}} (\beta_1 \alpha_2)^{m_{14}} (\beta_1 \beta_2 \alpha_3)^{m_{13}} (1 - \alpha_1 - \beta_1 \alpha_2 - \beta_1 \beta_2 \alpha_3)^{R_1 - r_1} \times \alpha_2^{m_{23}} (\beta_2 \alpha_3)^{m_{24}} (1 - \alpha_2 - \beta_2 \alpha_3)^{R_2 - r_2} \alpha_3^{m_{34}} (1 - \alpha_3)^{R_3 - r_3} \quad (11)$$

There are 5 parameters, in (11) and the non-identifiability of S_3 from p_4 is even clearer since β_3 does not appear.

As Davidson and Solomon (1974) showed, the MLEs are equivalent to MMEs based on minimal sufficient statistics of an exponential family distribution, so long as the dimension of the minimal sufficient statistics equals to dimension of the parameter space. The algorithm for solving MMEs is oftentimes much simpler than solving the system of equations resulting from differentiating the log-likelihood with respect to each parameter (see Brownie, et al. (1985), pages 170-175 for more discussion of this issue). The minimal sufficient statistics can be found in the $J_1 F_1 A_1 . I$ model by collecting together the exponents for each of the five parameters and showing that these summed exponents are the only statistics in the likelihood.

Let

$$\begin{aligned}m_1 &= m_{12} \\ m_2 &= m_{13} + m_{14} \\ m_3 &= m_{13} + m_{23} \\ m_4 &= m_{14} + m_{24} \\ m_5 &= m_{34}\end{aligned}$$

Therefore $r_1 = m_1 + m_2$, $r_2 = m_3 + m_4 - m_2$, and $r_3 = m_5$.

Re-express (11):

$$L \propto \alpha_1^{m_1} \beta_1^{m_2} \alpha_2^{m_3} \beta_2^{m_4} \alpha_3^{m_5} (1 - \alpha_1 - \beta_1 \alpha_2 - \beta_1 \beta_2 \alpha_3)^{R_1 - (m_1 + m_2)} (1 - \alpha_2 - \beta_2 \alpha_3)^{R_2 - (m_3 + m_4 - m_2)} (1 - \alpha_3)^{R_3 - m_5} \quad (12)$$

The likelihood (12) meets the requirements of Davidson, and Solomon (1974), thus setting $m_i = E(m_i)$ yields a system of equations whose solution is identical to the MLEs.

$$\begin{aligned}m_1 &= R_1 \alpha_1 \\ m_2 &= R_1 \beta_1 \alpha_2 + R_1 \beta_1 \beta_2 \alpha_3 \\ m_3 &= R_1 \beta_1 \alpha_2 + R_2 \alpha_2 \\ m_4 &= R_1 \beta_1 \beta_2 \alpha_3 + R_2 \beta_2 \alpha_3 \\ m_5 &= R_3 \alpha_3\end{aligned}$$

To solve the above system of equations, first re-express each of the equations in terms of a single parameter. The approach taken here is to re-express with the sequence $\alpha_1, \beta_1, \alpha_2, \beta_2$, and as matched with m_1, m_2, m_3, m_4 , and m_5 . (Hats over the parameter estimates are suppressed throughout to minimize notation.)

$$\begin{aligned}\alpha_1 &= \frac{m_1}{R_1} \\ \beta_1 &= \frac{m_2}{R_1\alpha_2 + R_1\beta_2\alpha_3} \\ \alpha_2 &= \frac{m_3}{R_1\beta_1 + R_2} \\ \beta_2 &= \frac{m_4}{R_1\beta_1\alpha_3 + R_2\alpha_3} \\ \alpha_3 &= \frac{m_5}{R_3}\end{aligned}$$

Now work from the last equation backwards re-expressing the equations in terms of expressions evaluated below them.

$$\begin{aligned}\alpha_3 &= \frac{m_5}{R_3} \\ \beta_2 &= \frac{m_4 R_3}{R_1\beta_1 m_5 + R_2 m_5} \\ \alpha_2 &= \frac{m_3}{R_1\beta_1 + R_2} \\ \beta_1 &= \frac{R_2 m_2}{R_1 r_2} \\ \alpha_1 &= \frac{m_1}{R_1}\end{aligned}$$

The value for β_1 is substituted into the remaining two equations with parameters remaining on the right hand side.

$$\begin{aligned}\alpha_1 &= \frac{m_1}{R_1} \\ \beta_1 &= \frac{R_2 m_2}{R_1 r_2} \\ \alpha_2 &= \frac{r_2 m_3}{R_2(m_3 + m_4)} \\ \beta_2 &= \frac{m_4 R_3 r_2}{R_2 m_5(m_3 + m_4)} \\ \alpha_3 &= \frac{m_5}{R_3}\end{aligned}$$

Re-expressing in terms of S and p and using the original m_{ij} ,

$$S_1 = \frac{r_2 m_{12} + R_2(r_1 - m_{12})}{R_1 r_2} \quad (13)$$

$$p_2 = \frac{r_2 m_{12}}{r_2 m_{12} + R_2(r_1 - m_{12})} \quad (14)$$

$$S_2 = \frac{r_2[(m_{13} + m_{23})m_{34} + R_3(m_{14} + m_{24})]}{R_2 m_{34}(m_{13} + m_{14} + r_2)} \quad (15)$$

$$p_3 = \frac{(m_{13} + m_{23})m_{34}}{(m_{13} + m_{23})m_{34} + R_3(m_{14} + m_{24})} \quad (16)$$

$$S_3 p_4 = \frac{m_{34}}{R_3}$$

MLEs for $J_i F_j A_k . I$

The same approach used for $J_1 F_1 A_1 . I$ can be used to find the MLEs for $J_i F_j A_k . I$, where $i, j, k > 0$. Let $h = i + j + k$. So long as there are releases at each point of recapture all survival and tag recapture rates can be estimated except for the last period. I.e., $2(h-1)+1$ parameters can be estimated, where the additional parameter is $S_h p_{h+1}$. The algorithm is sketched here.

Again reparametrize S_i and p_{i+1} in terms of α_i and β_i . All the exponents in the likelihood for each parameter except α_h are summed together as before. These sums and $m_{h-1,h}$ form a minimal sufficient statistic set for the parameters $\alpha_1, \dots, \alpha_h, \beta_1, \dots, \beta_{h-1}$. Let the notation m_1 denote the sum of exponents corresponding to α_1, m_2 for β_1, m_3 for α_2 , and so on. The estimates are found by solving for β_{h-1} , then plugging in α_h , then going back to β_{h-2} and using estimates of β_{h-1} and α_{h-1} and α_h . One then moves back and forth through the parameter list, substituting *later* parameters in solutions for earlier parameters.

Fortunately a pattern exists in the solutions for arbitrary h .

$$\begin{aligned}\alpha_i &= \frac{r_i m_{2i-1}}{R_i(m_{2(i-1)} + r_i)} \quad i = 2, \dots, h-1 \\ \alpha_h &= \frac{r_h}{R_h} \\ p_i &= \frac{R_{i+1} r_i m_{2i}}{R_i r_{i+1} (m_{2(i-1)} + r_i)} \quad i = 2, \dots, h-1,\end{aligned}$$

where $m_0 = 0$.

The estimates for S_i and p_{i+1} are then:

$$\begin{aligned}S_i &= \alpha_i + \beta_i \\ &= \frac{r_i}{R_i(m_{2(i-1)} + r_i)} \left[m_{2i-1} + \frac{R_{i+1} m_{2i}}{r_{i+1}} \right] \quad (17)\end{aligned}$$

$$\begin{aligned}p_{i+1} &= \frac{\alpha_i}{\alpha_i + \beta_i} \\ &= \frac{m_{2i-1}}{m_{2i-1} + \frac{R_{i+1} m_{2i}}{r_{i+1}}} \quad (18)\end{aligned}$$

The mles were originally derived by Cormack (1964). Burnham, et al. (-1987, p. 114) give the estimators using m-array notation. I have not seen a method of moments solution presented in this manner before, however. Below I equate the Burnham, et al. formulation with the one given here. First some additional notation used by Burnham, et al. (with minor changes):

$$\begin{aligned}m_j^* &= \sum_{i=1} m_{ij} \\ z_j &= \sum_{i=1}^{j-1} \sum_{n=j+1}^h m_{in}\end{aligned}$$

m_j^* is then the j th recapture site column total, which for my notation equals the minimal sufficient statistics matched with the $\alpha_j, j = 1, \dots, h-1$. So Burnham, et al.'s m_j^* equals my m_{2j-1} (all the odd numbered j) excluding that corresponding to α_h . The z_j correspond to my m_{2j} (the even numbers), or the statistics corresponding to β_j . The estimators given by Burnham, et al. in their notation are:

$$\begin{aligned}S_i &= \frac{r_i}{R_i} \left[\frac{m_{i+1}^*}{z_i + r_i} + \frac{z_{i+1} R_{i+1}}{(z_i + r_i) r_{i+1}} \right], \quad i = 2, \dots, h-1 \\ p_{i+1} &= \frac{m_{i+1}^*}{m_{i+1}^* + z_{i+1} R_{i+1} / r_{i+1}}, \quad i = 1, \dots, h-1\end{aligned}$$

Substituting z and m^* for 'the m in Equations (17) and (18) yields identical results.

Variance estimates

Asymptotic variances of \hat{S}_i and \hat{p} can be found by using Taylor series approximations for the estimators (the Delta method). The derivations involve a lot of algebra and are omitted. The resulting estimates of the variances are from Burnham, et al. (1987, p. 115).

$$\begin{aligned}\hat{V}(\hat{S}_i) = & \hat{S}_i^2 [1/r_i - 1/R_i + \hat{q}_{i+1}^2 (1/r_{i+1} - 1/R_{i+1}) \\ & + \hat{q}_{i+1}^2 \left((1 - r_{i+1}/R_{i+1})^2 \frac{m_{i+1}^*}{z_{i+1}(z_i + r_i)} \right)]\end{aligned}\quad (19)$$

$$\hat{V}(\hat{p}_{i+1}) = (p_{i+1}q_{i+1})^2 [1/r_{i+1} - 1/R_{i+1} + 1/m_{i+1}^* + 1/z_{i+1}] \quad (20)$$

B MLEs for Protocol II

Using the same reparametrization of S_i and p_{i+1} to α_i and β_i as for Protocol I, the Protocol II likelihood for the case $i = j = k = 1$ is:

$$\begin{aligned}L \propto & \alpha_1^{m_{12}} (\beta_1 \alpha_2)^{m_{13}} (\beta_1 \beta_2 \alpha_3)^{m_{14}} (1 - \alpha_1 - \beta_1 \alpha_2 - \beta_1 \beta_2 \alpha_3)^{R_1 - r_1} \times \\ & \alpha_2^{m_{23}} (\beta_2 \alpha_3)^{m_{24}} (1 - \alpha_2 - \beta_2 \alpha_3)^{R_2 - r_2}\end{aligned}\quad (21)$$

Note that β_2 and α_3 are inseparable, hence S_2 and p_3 cannot be estimated, since $S_2 = \alpha_2 + \beta_2$; θ is substituted for $\beta_2 \alpha_3$.

As in Protocol I the method of moments estimator based on the minimal sufficient statistics equals the MLE. Using the same technique as for Protocol I, the minimal sufficient statistics are $m_1 = m_{12}$, $m_2 = m_{13} + m_{14}$, $m_3 = m_{13} + m_{23}$, and $m_4 = m_{14} + m_{24}$. The method of moments system of equations:

$$\begin{aligned}m_1 &= R_1 \alpha_1 \\ m_2 &= R_1 \beta_1 \alpha_2 + R_1 \beta_1 \theta \\ m_3 &= R_1 \beta_1 \alpha_2 + R_2 \alpha_2 \\ m_4 &= R_1 \beta_1 \theta + R_2 \theta\end{aligned}$$

The maximum likelihood estimates (in terms of S_i and p_{i+1}) are:

$$\begin{aligned}S_1 &= \frac{(r_1 - m_{12})R_2 + r_2 m_{12}}{R_1 r_2} \\ Pa &= \frac{m_{12} r_2}{m_{12} r_2 + (r_1 - m_{12})R_2} \\ S_2 p_3 &= \frac{r_2}{R_2} \left(\frac{m_{13} + m_{23}}{m_{13} + m_{14} + m_{23} + m_{24}} \right) \\ S_2 q_3 S_3 p_4 &= \frac{r_2}{R_2} \left(\frac{m_{14} + m_{24}}{m_{13} + m_{14} + m_{23} + m_{24}} \right)\end{aligned}$$

The estimators for S_1 and p_2 are identical to those under Protocol I (equations (13) and (14)). This will be true for all survival and recapture rates up to the last juvenile dam. The variance estimates for these parameters are therefore the same as for Protocol I. It is the parameters just, prior to and after the fishery that become compounded with each other under Protocol II.

Some extensions

I consider briefly a few special cases of Protocol II, namely the situations with more than one Juvenile dam ($i \geq 1$) and/or more than one Adult dam ($k \geq 1$).

The first extension is $J_2F_1A_1.II$, an additional juvenile dam. The parameters in terms of α_i and β_i are $\alpha_1, \dots, \alpha_4$ and β_1, \dots, β_3 . $\beta_3\alpha_4$ are inseparable, therefore, the mortality prior to the fishery cannot be estimated ($S_3 = \alpha_3 \cdot t \beta_3$). The estimates of $\alpha_1, \dots, \alpha_3$ and β_1 and β_2 are identical to those in Protocol I. Therefore estimates of S_1, S_2, p_2, p_3 are those in equations (17) and (18). The estimate of the compound parameter, $\beta_3\alpha_4$ or $S_3p_4S_4p_5$,

$$S_3p_4S_4p_5 = \frac{r_3(m_{15} + m_{25} + m_{35})}{R_3(r_3 + m_{14} + m_{15} + m_{24} + m_{25})}$$

A second extension is $J_1F_1A_2.II$, an additional adult dam. Only S_1, p_2 , and p_4 (the first adult dam recapture rate) are separately estimable. The estimates for S_1 and p_2 are the same as in (13) and (14).

$$p_4 = \frac{\beta_2\alpha_3}{\beta_2\alpha_3 + \beta_2\beta_3}$$

where $\beta_2\alpha_3$ and $\beta_2\beta_3$ are inseparable.

A third extension is the case with two juvenile dams and two adult dams, $J_2F_1A_2.II$. The estimable parameters are S_1, p_2, S_2, p_3 , and p_5 , i.e., survival prior to the two juvenile dams, and recapture at the two juvenile dams and the first adult dam.

C MLEs for Protocol III

The data for $J_1F_1A_1.III$ was summarized in the m-array in Table 1. The likelihood for the observed data is complicated because of a summation over combinations of m_{13} and m_{23} that add to m_3 . To minimize notation, let p_{ij} denote the conditional probability of 'first time' recovery at site j, given release at site i, with the special case that 'site' 5 is not being recaptured anywhere.

$$L = \sum_k \binom{R_1}{m_{12}, k, m_{14}} p_{12}^{m_{12}-k} p_{13}^k p_{14}^{m_{14}-k} p_{15}^{R_1-k} \binom{R_2}{m_3-k, m_{24}} p_{23}^{m_3-k} p_{24}^k p_{25}^{R_2-(m_3-k)},$$

where $R_1^* = R_1 - m_{12} - m_{14}$ and $R_2^* = R_2 - m_{24}$, and the lower and upper limits on summation are $\max(0, m_3 - R_2^*)$ and $\min(R_1^*, m_3)$, respectively. Calculating this likelihood to determine MLEs then involves some messy summation, but may not be that computationally difficult for realistic values of m_3 .

An alternative approach to finding MLEs is the EM algorithm. If the conditional-expectation of m_{13} (and consequently m_{23}) is substituted for the unobserved value m_{13} , then the m-array looks exactly like the Protocol II model which has closed form solutions. The E-step in this problem is to calculate the conditional expectation of m_{13} given m_3 , the rest of the observed data, and estimates of the parameters. Unfortunately calculating the conditional expectation involves an awkward summation as well; The conditional probability (letting θ represent the vector of parameters),

$$\Pr(m_{13} = X | m_{12}, m_3, m_{14}, R_1, m_{24}, R_2, \theta) = \frac{\binom{R_1^*}{X} p_{13}^X p_{15}^{R_1^*-X} \binom{R_2^*}{m_3-X} p_{23}^{m_3-X} p_{25}^{R_2^*-(m_3-X)}}{\sum_h \binom{R_1^*}{h} p_{13}^h p_{15}^{R_1^*-h} \binom{R_2^*}{m_3-h} p_{23}^{m_3-h} p_{25}^{R_2^*-(m_3-h)}}$$

An approximation to the conditional expectation is 'to use the expected recoveries conditioning only within each release group at first ignoring the constraint m_3 . In other words calculate

$$a = E \left[m_{13} | R_1^*, \frac{p_{13}}{1 - p_{12} - p_{14}} \right] = R_1^* \frac{p_{13}}{1 - p_{12} - p_{14}}$$

$$b = E \left[m_{23} | R_2^*, \frac{p_{23}}{1 - p_{24}} \right] = R_2^* \frac{p_{23}}{1 - p_{24}}$$

Then approximate $E(m_{13}|m_{12}, m_{13}, m_{14}, R_1, m_{24}, R_2, \theta)$ with

$$\hat{m}_{13} = m_{13} \frac{a}{a+b}.$$

I have not proven this, but believe that this approximation is asymptotically equivalent to using the conditional expectations. The performance with simulated data sets has been quite good, in any case. Further study of the properties of 'this estimator is needed.

Adding additional adult recapture sites, say $J_1 F_1 A_k .III$, where $k > 1$, turns out to be no different. The fishery recoveries are partitioned into m_{13} and m_{23} in the same manner. Given this imputed m-array, the (approximate) EM algorithm can begin. Adding additional juvenile recapture sites, say $J_i F_1 A_k .III$, where $i > 1$ and $k \geq 1$, requires partitioning the fishery recoveries, m_{13} , into i subgroups, but the approximate E step, discussed previously, is extended in an obvious way. Extensions to j fisheries have not been considered.

For variance estimates for EM-based maximum likelihood estimates see Tanner (1996).

D Likelihoods for mixed tagging

For each of the three cases of mixed tagging considered, the likelihood is a product of two or three likelihoods, one always being the likelihood for the subgroup receiving both tags. Depending upon the release-recapture protocol, the likelihood components will vary.

Case 1: Combination and CWT-only subgroups

The likelihood is a product of the likelihood for the combined group and a binomial likelihood for the CWT-only recoveries. E.g., under Protocol I, with $i = j = k = 1$, the full likelihood is the product of equation (1) and

$$L(\text{CWT}) \propto (S_1 S_2 p_3)^{m_{13C}} (1 - S_1 S_2 p_3)^{R_C - m_{13C}}$$

R_C denotes the release size of CWT-only fish and m_{13C} the fishery recoveries. Iterative methods are required to estimate the MLEs of the estimable parameters based on these combined likelihoods.

Case 2: Combination and PIT-only subgroups

The EM algorithm lends itself well to the calculation of MLEs in this case. The unobserved recoveries in the fishery are easily imputed, because the conditional distribution for m_{13} is binomial. E.g., m_{13} given the other cells is $\text{binomial}(R_1 - m_{12} - m_{14}, \frac{\pi_{13}}{1 - \pi_{12} - \pi_{14}})$ and similarly $m_{23} \sim \text{binomial}(R_2 - m_{24}, \frac{\pi_{23}}{1 - \pi_{24}})$. So the imputed values are the expectations for these binomials.

Under Protocol I, the likelihood component for the PIT only group including imputed-values is, for $i = j = k = 1$,

$$\begin{aligned} L(\text{PIT}) \propto & (S_1 p_2)^{m_{12P}} (S_1 q_2 S_2 p_3)^{\hat{m}_{13P}} (S_1 q_2 S_2 q_3 S_3 p_4)^{m_{14P}} \\ & (1 - S_1 p_2 - S_1 q_2 S_2 p_3 - S_1 q_2 S_2 q_3 S_3 p_4)^{R_{1P} - m_{12P} - \hat{m}_{13P} - m_{14P}} \\ & (S_2 p_3)^{\hat{m}_{23P}} (S_2 q_3 S_3 p_4)^{m_{24P}} (1 - S_2 p_3 - S_2 q_3 S_3 p_4)^{R_{2P} - \hat{m}_{23P} - m_{24P}}, \end{aligned}$$

where the subscript P refers to the PIT-only group. $L(\text{PIT})$ is identical to the combination data's likelihood except for the last two terms representing releases from the fishery. Exact closed form solutions may exist, but have not been worked out, yet. Iterative solutions are relatively simple, however.

Under Protocol II, the parameter grouping is identical for both components of the likelihood. The exponents are simply combined, e.g., for $S_1 q_2 S_2 p_3$ the exponent is $m_{13} + \hat{m}_{13P}$.

Case 3: Combination, PIT-only, and CWT-only subgroups

The likelihood under any protocol is simply the product of three likelihoods, one for each subgroup:

$$L = L(\text{combination})L(\text{PIT})L(\text{CWT})$$

The Design and Analysis of Salmonid Tagging Studies in the Columbia Basin

Other related publications, reports and papers not included in this series **available** through, the professional literature or from the Bonneville Power Administration (BPA) Public Information Center - **CKPS-1**, P.O. Box 362 1, **Portland**, OR 97208.

1997

Newman, K. 1997. Bayesian averaging of generalized **linear** models for passive integrated tag recoveries from salmonids in the Snake River. North American Journal of Fisheries Management 17: 362-377.

1996

Skalski, J. R. 1996. Regression of abundance estimates from mark-recapture surveys against environmental variables. Can. J. Fish. Aquat. Sci. 53: 196-204.

Skalski, J. R., R. L. Townsend, R. F. Donnelly, and R. W. Hilborn. 1996. The relationship between survival of Columbia River fall chinook salmon and **inriver** environmental factors: Analysis of historic data for juvenile and adult **salmonid** production. Final Report Phase **II**. Technical Report (DOE/BP-35885-10) to BPA.

Smith, S. G., J. R. Skalski, J. R., J. W. Schlechte, A. Hoffmann and V. Cassen. 1996. Introduction **SURPH.1** analysis of release-recapture data for survival studies. Technical Report **DOE/BP-02341-3**) to BPA.

1995

Newman, K. 1995. Adult **salmonid PIT-Tag** returns to Columbia River's Lower Granite Dam. **Technical Report** (D&BP-35885-5) to BPA.

1994

Smith, S. G., J. R. Skalski, J. R., J. W. Schlechte, A. Hoffmann and V. Cassen. 1994. **SURPH.1** Manual: Statistical **survival** analysis of fish and wildlife tagging studies. Technical Report (DOE/BP-02341-2) to BPA.

1993

Dauble, D. D., J. R. Skalski, A. Hoffmann, and A. E. Giorgi. 1993. Evaluation and application of statistical methods for estimating smolt survival. Technical Report (**DOE/BP-62611-1**) to BPA.

Skalski, J. R., A. Hoffmann and S. G. Smith. 1993. Development of survival relationships using concomitant variables measured from individual smolt implanted with PIT-tags. Annual Report **1990-1991** (DOE/BP-02341-I) to BPA.

Skalski, J. R., and A. E. Giorgi. 1993. Juvenile Passage Program: A plan for estimating **smolt travel** time and survival in the Snake and Columbia Rivers. **Technical Report (DOE/BP-35885-3)** to BPA.

Smith, S. G., J. R. Skalski, and A. E. Giorgi. 1993. Statistical evaluation of travel time estimation based on data from freeze-branded chinook salmon on the Snake River, 1982-1990. Technical Report (**DOE/BP-35885-4**) to BPA.

1991

Giorgi, A. E. 1990. Mortality of yearling chinook salmon prior to arrival at Lower Granite Dam on the Snake River. Technical Report (DOE/BP-16570-1) to BPA.

Bonneville Power Administration
PO Box 3621 Portland, Oregon 97208-3621

DOE/BP-35885-11 APRIL 1998 601

